DOI: 10.35191/medsoft_2020_1_32_5_8

BEZPEČNOST DAT A BIG DATA V BIOMEDICÍNĚ

Jiří Berger, Jiří Kofránek

Abstrakt

V dnešní době je patrný trend digitalizace zdravotnických archivů a související dokumentace, nastává tedy čas na zapojení technologií označovaných Big Data v oblasti biomedicínské informatiky. Tyto technologie nabízí rychlejší a efektivnější zpracování a sdílení obrovského množství dat. Vzhledem k tomu, že zdravotní péče pracuje s velmi citlivými daty, je jedním z hlavních zájmů ochrana dat pacientů. V mnoha zemích probíhá programové zavádění elektronizace zdravotní péče. Například v USA probíhá „Health Information Technology for Economic and Clinical Health Act", (HITECH). Cílem výzkumu je návrh a definice pravidel, která zamezí zneužití a únikům citlivých biomedicínských dat. Současně však v minimální míře omezí efektivitu jejich zpracování a kvalitu výstupních dat. Hromadnost zpracování osobních a citlivých dat se postupně stává obrovským rizikem a současně příležitostí pro nastavení pravidel a procesů vedoucí k minimalizaci, či dokonce eliminaci těchto rizik.

Big Data skrývají obrovský potenciál pro výzkum v oblasti biomedicíny v mnoha oblastech ať již při analýze segmentace pacientů, cen a výsledků léčby, kde umožní zjistit zdravotně a cenově nejefektivnější postup léčení pro konkrétního pacienta a také například v proaktivní identifikaci pacientů, u nichž by se vyplatila zdravotnická prevence. Principiálně jsou Big Data použitelná k tomu, aby z analýzy výskytu chorob bylo možné dělat epidemiologické závěry a navrhovat preventivní opatření, mohou pomáhat při detekci a minimalizaci pokusů o podvody ve zdravotnictví a veřejném zdravotním pojištění a také přinášejí příležitost spolupráce s farmaceutickými společnostmi tak, aby pro ně bylo snazší identifikovat skupinu relevantních pacientů pro klinické testy (za předpokladu předchozího souhlasu pacientů a dodržení etických norem).

Článek si bere za cíl vysvětlit rozdílné přístupy a oblasti bezpečnosti v souvislosti s hromadným zpracováním dat ať již z pohledu bezpečnosti databázových dat jako celku, zamezení možnosti nepřímého získávání konkrétních údajů z neanonymizovaného nebo i anonymizovaného souboru dat a z oblasti netriviálního dotazování.

Úvod

Čím větší množství heterogenních biomedicínských informací se podaří sdružit pod technologie Big Data tak, aby obsahovaly co nejkompletnější záznamy, tím vyšší hodnotu mohou mít pro budoucí zpracování různých druhů analýz na základě zdravotní dokumentace celé populace a souvisejících biomedicínských informací.

Jakmile dojde k nastavení takového projektu, ač autoři tohoto článku připouští, že větší než technický to bude organizačně etický problém, bude nejefektivnějším způsobem práce s takovým množstvím dat zpřístupnění anonymizovaných informací odborné veřejnosti jako zdroj výzkumu. V zahraničí je obvyklé, že pokud je projekt zcela nebo i jen částečně financován a podporován z veřejných zdrojů, bývají nastaveny podmínky tak, aby byly informace plně dostupné k nekomerčním aktivitám s minimálními omezeními. I přesto, že biomedicínské informace, obzvláště s ohledem na citlivost uchovávaných dat, nebude nikdy možné poskytovat zcela bez omezení, stále existuje široká škála použití, implementací a aplikací, které by takovým přístupem disponovaly.

To však, obzvláště díky použití populačních dat, s sebou přináší rizika možnosti nepřímé identifikace konkrétních informací o pacientech. Sebemenší náznak zneužití přináší etické problémy a současně může zastavit veškerý související výzkum.

Aby se předešlo možnému zneužití, je nutné nastavit velmi přísná a efektivní pravidla umožňující maximální výtěžnost dat při současném striktním zachování anonymity a ochrany osobních údajů. Současně je nutné regulovat způsoby vytěžování dat tak, aby nemohlo dojít ke zneužití nebo úniku citlivých informací jakýmkoliv i jen teoreticky proveditelným způsobem. V zahraničí již existuje odpovídající legislativa, např. aktuální verze „Health Insurance Portability and Accountability Act" (HIPAA) v USA která specifikuje standardy transakcí se zdravotními záznamy, obdobně směrnice EU Data Protection Directive 95/46/EC, která definuje požadavek souhlasu pacienta se zpracováním jeho údajů a přenositelnost dat. EU však stále nemá jednotný přístup k ochraně citlivých údajů [1]. Problematika se tak rozděluje mezi vlastní metody zabezpečení pomocí šifrování a dále se zaměřuje na specifika, která pro oblast zabezpečení znamená využití Big Data.

Šifrování v biomedicíně a zdravotnictví

Vlastní šifrování biomedicínských a zdravotnických dat různými metodami pak přináší své výhody i nevýhody a zatím neexistuje jednotná metoda, která by byla vhodná pro všechny aplikace. V souhrnu bychom rádi představili alespoň ty nejznámější včetně principu jejich fungování, výhod a nevýhod.

Základní ochrana

Základní ochranu lze popsat jako množinu úrovní přístupových práv, která jsou nastavena nad daty tak, aby každý uživatel měl přístup k požadované podmnožině dat, kterou může plně používat – prohledávat, analyzovat apod. Tento princip nejblíže odpovídá nastavení práv například na operačním systému UNIX. Superuživatel vidí všechna data, standardní uživatelé vidí jen to, co jim dovoluje úroveň jejich přístupu. Výhodou je možnost provádět jakékoliv výpočty a vyhledávání za pomocí aplikací, která mají vyšší práva než standardní uživatel a mohou definovat úroveň zobecnění výsledku do takové míry, aby standardní uživatel získal jen obecná, generalizovaná, agregovaná nebo odpovídajícím způsobem anonymizovaná data. Nevýhodou tohoto přístupu je neexistující ochrana proti vlastníkovi platformy. Tento způsob je využíván při zpracování hromadných dat například u platformy Hadoop díky Hadoop Distributed File System (HDFS).

Dvouvrstevné šifrování

Dvouvrstevná architektura šifrování je zaměřena oproti výše uvedenému základnímu šifrování na ochranu proti vlastníkovi platformy. Data jsou šifrovaná tzv. end-to-end. Z toho plyne, že i kdyby vlastník nebo provozovatel platformy chtěl k datům přistupovat, nemá k tomu prostředky, jelikož data jsou šifrovaná klíčem, kterým nedisponuje. To lze považovat za výhodu, ale oproti tomu nevýhodou je nemožnost vyhledávat v datech. Principiálně je tento způsob nastaven tak, že existují dva klíče, z nichž jeden drží Key Management Server (KMS), obsluhovaný majitelem nebo provozovatelem platformy, druhý drží uživatel. Jeden bez druhého tedy nemůže k datům přistupovat.

Vícevrstevné šifrování (Multi party computation)

Vícevrstevné šifrování je podobné principu důvěryhodného subjektu, kterému jednotliví uživatelé důvěřují a předávají svoje vstupy. Tento subjekt počítá výstupy s využitím definovaného algoritmu. Takový způsob šifrování se ale plně objede bez důvěryhodného subjektu. Jednotlivé servery si vyměňují data zašifrovaná vícevrstevným šifrováním, avšak každý z nich má přístup jen ke svým vstupům a výstupům. Servery společně pouze díky posílání zašifrovaných zpráv spočítají vý-

DOI: 10.35191/medsoft_2020_1_32_5_8

stup místo důvěryhodného subjektu. Tento způsob je možno využít v případě, že máme k dispozici více serverů, ale hrozí riziko, že některé z nich budou kompromitovány. Algoritmus je možno konfigurovat pomocí parametrů *t* a *n* tak, aby kompromitace maximálního počtu *t* serverů z celkového počtu *n* serverů nezpůsobila žádný únik dat.

Plně homomorfní šifrování (Fully homomorphic encryption)

Metoda plně homomorfního šifrování (FHE) je založena na principu, kdy databáze obsahuje kompletně zašifrovaná data, ke kterým se nemůže majitel nebo provozovatel platformy dostat. Uživatelé se mohou k databázi připojit a zadávat jí dotazy, úkoly, vyhledávat v ní, nebo případně zadávat výpočty. Databáze takové zadání přijme a úkol provede. Jako výsledek vrátí data, která jsou uživateli srozumitelná, ale samotná databáze a tedy i vlastník nebo provozovatel platformy nemá možnost takovému výpočtu, případně uloženým datům porozumět, nebo k nim přistupovat. Jde o jeden z pokročilých principů šifrování, který je zatím ve vývoji. V současnosti existuje několik pilotních a výzkumných projektů, které se snaží plně homomorfní šifrování implementovat, ale zatím ještě není možno jej nasadit do praktického provozu. Z pohledu procesů a požadavků na bezpečnost při práci s biomedicínskými daty jde o jednu z cest, která by se v budoucnosti mohla svými vlastnostmi ukázat jako jedna z vhodných platforem. Autoři tohoto textu se domnívají, že v mnoha konkrétních úlohách nad daty pacientů je tento způsob vyhledávání preferovaným způsobem, u nějž převažují klady nad zápory.

Částečně homomorfní šifrování (Somewhat homomorphic encryption)

Částečně homomorfní šifrování (SHE) je ve svém principu způsob šifrování podobný výše popsanému přístupu FHE. Stejně jako v případě FHE, provádí databáze výpočty nad daty, kterým nijak nerozumí. Hlavní odlišností od FHE je nemožnost v databázi, která používá částečně homomorfní šifrování, vyhledávat. Typickou úlohou pro SHE jsou různé statistické analýzy a výstupy a proto je ideálním nástrojem, který chrání konkrétní data jak před vlastníkem nebo provozovatelem platformy, tak před koncovým uživatelem. Ač se tento přístup může zdát méně hodnotný než FHE, existuje řada úloh, pro které může být tento způsob šifrování výhodný, nebo dokonce žádoucí z pohledu zadání a z pohledu procesního nastavení bezpečnostních pravidel. Autoři tohoto textu se domnívají, že pro potřeby statistických výstupů patří tento způsob šifrování k nejbezpečnějším z pohledu ochrany uložených dat.

Šifrování řízené pacientem (Patient controlled encryption)

Z pohledu pacientských dat (nikoliv však obecně biomedicínských) se v poslední době jako vhodný způsob zabezpečení a šifrování prosazuje šifrování řízené pacientem (PCE). Základní premisou je na rozdíl od předchozích přístupů kontrola nad šifrováním dat na straně pacienta. Data mají hierarchickou strukturu, a pacient určuje, které uzly této struktury mohou číst kteří lékaři. Z pohledu běžných životních situací se tento přístup může zdát efektivní a navíc i etický, jelikož sám pacient si rozhoduje o tom, která data poskytuje. Na druhou stranu však v mnoha případech neznalost na první pohled nesouvisejících informací (např. o úrazu, operaci atd.) může vést lékaře ke stanovení odlišné diagnózy, než by určil v případě, že by rozhodoval v plném kontextu. Dalším z problematických bodů je urgentní péče o pacienta v okamžiku, kdy není schopen z pohledu svého akutního stavu rozhodovat o poskytnutí dostatečného přístupu pro lékaře.

Šifrování za pomocí blockchain

V posledních letech probíhá intenzivní výzkum v oblasti blockchain. Na rozdíl od výše popisovaných principů nejde z pohledu technologie blockchain přímo o šifrování dat, ale spíše o formu zabezpečení přístupu k nim, případně průkaznou identifikaci o přístupu k definovaným datům. Blockchain je v těchto případech používán převážně k evidenci toho kdo, kdy a k jakým konkrétním datům přistupoval. V kombinaci s anonymizací dat pak vzniká nástroj, který na opt-in nebo opt-out principu zajistí přístup k datům a na základě ověření tohoto přístupu pak lze následně přistoupit k deanonymizaci. Vzhledem k tomu, že tato oblast je relativně mladá a je atraktivní pro různé startupové projekty, jde z velké části o hledání způsobu, jak by bylo možné blockchain využít k zabezpečení biomedicínských (nebo jakýchkoliv jiných) dat než o skutečné implementace navržených postupů. Autoři nevylučují, že některé z těchto cest se ukážou samostatně nebo v kombinaci s některým z výše uvedených způsobů šifrování jako životaschopné pro konkrétní implementace, zatím k tomu však nedošlo. Z podstaty této technologie a principů s ní spojených lze základní premisy výzkumných projektů popsat tak, že existují databáze (ve své podstatě nezávisí na tom, zda přímo v blockchainu nebo mimo něj), které obsahují biomedicínská data. Vedle toho existují uživatelé těchto dat. Mezi uživateli a databázemi je blockchain, který drží informace o tom, jaké dotazy byly položeny, a jaké odpovědi byly na každý dotaz vráceny. A z podstaty blockchainu je díky distribuci klíčů tato informace neměnná (a to ať úmyslně, nebo nahodilým vlivem). Dalšími možnými využitími pak může být například kontrola verzování dat v databázích včetně identifikace původce změn každé verze.

Uplatnění Big Data v biomedicíně a zdravotnictví

Využívání Big Data v biomedicíně a zdravotnictví bude mít vždy svá specifika. Míra anonymizace použitých dat bude nepřímo úměrná kvalitě výstupů [2]. Znamená to, že jedním z klíčových prvků úspěšného využití Big Data v biomedicíně a zdravotnictví bude nastavení hranice mezi anonymizací a způsobem vytěžování dat. Pro efektivní využití by bylo potřebné provést zcela základní anonymizaci, která z dat odstraní (nebo zaručeným způsobem znepřístupní) jen základní osobní informace jako je jméno, příjmení a rodné číslo a nahradí je anonymním údajem, který však přesto zajistí identifikaci subjektu napříč daty. Takto upravená data budou vždy zranitelná, a proto je nutné jednotlivým způsobům potenciálních útoků efektivně předcházet.

Například dotaz, který vrací předepsaná léčiva a jejich dávkování konkrétnímu pacientovi obsahuje citlivá data. Ze znalosti druhů léčiv lze usuzovat diagnózu pacienta. Pokud budou osobní údaje anonymizované, lze usoudit, že poskytnutá data nebudou citlivá.

Oproti tomu typickými dotazy, které poskytují data, jenž nejsou citlivá, jsou například: Dotaz který získá počet pacientů daného lékaře, nemusí nutně obsahovat citlivá data. Stejně tak, dotaz na seznam předepsaných léčiv v daném regionu nevrací citlivé výsledky. Podobně výskyt specifické diagnózy napříč populací neposkytuje citlivé údaje.

Bezpečnost databázových dat jako celku

Jednou z metod zvýšení bezpečnosti dat v biomedicíně a lékařství je šifrování podkladových dat, které přidává další bezpečnostní vrstvu navíc, a právě návrh této architektury umožňuje významně snížit možnosti zneužití dat [3].

Existují specializované pokročilé algoritmy [4] umožňující šifrování medicínských záznamů tak, že umožní jejich rozšifrování pouze osobám majícím relevantní oprávnění. Algoritmy mají několik výhod oproti klasickým šifrovacím způsobům (symet-

DOI: 10.35191/medsoft_2020_1_32_5_8

rické i asymetrické šifry) – oproti RSA konceptu jsou rychlejší, levnější a flexibilnější, oproti symetrickým šifrám poskytují bezpečnost v případě vyzrazení sdíleného hesla. V [5] popisují autoři možnosti vyhledávání nad takto šifrovanými medicínskými daty. Tento algoritmus je ideální pro koncept Big Data.

Šifrování přinese zpomalení vyhledávání, proto jej nelze doporučit pro použití pro celé úložiště, nýbrž pouze pro základní data pacientů, u kterých je jasně daná struktura.

Nepřímé získání konkrétních dat z neanonymizovaného souboru dat

V případě, že se v databázi nacházejí kompletní neanonymizovaná data, nebo častěji data, která prošla pouze základní anonymizací, v rámci které byla nahrazena základní data pacientů (jméno, příjmení, rodné číslo), ale zbytek datového souboru je kompletní, je nutné detailně řešit principy omezení přístupu.

V takovém případě lze při cíleném útoku kombinací dotazů získat velmi konkrétní data, nebo přinejmenším data, která lze s vysokou pravděpodobností interpretovat zcela konkrétně.

Z výše uvedených důvodů je nutné zavést takové koncepční řešení, jenž spočívá v omezení dotazů, jejichž kombinace může odhalit citlivé údaje, případě takové kombinace umožnit jen osobám s vyšším oprávněním a současně zajistit zpětnou kontrolu a analýzu rizikovosti používaných dotazů a jejich výsledků.

Dalším rizikovým dotazem je takový, který ve svém výsledku předá významně malý počet výstupních entit.

Dotazy mohou obsahovat kombinaci několika faktorů. Pokud je ale daný dotaz „překombinován", může dojít v extrémním případě k situaci, že jeho výsledkem bude pouze jeden pacient, u kterého i bez znalosti jména můžeme odvodit, o koho se jedná. Například pokud známe i pouhou část chorobopisu daného člověka, dá se pomocí souvisejících informací (věk, pohlaví, bydliště) nepřímo získat jeho citlivé informace. Takové riziko se dá do určité míry eliminovat nasazením heuristických pravidel a jejich postupné zpřesňování, a tím zablokovat odpovědi, které by mohly obsahovat rizikovou množinu informací.

Nepřímé získání konkrétních dat z anonymizovaného souboru dat

V případě volby varianty anonymizace datového souboru lze za stanovených podmínek umožnit plný přístup do databáze.

Před uložením dat do databáze je možno (případně nutno, podle dané legislativy) data anonymizovat (odstranit jméno, rodné číslo) a zároveň generalizovat. Tomuto způsobu říkáme plná anonymizace. Generalizace spočívá ve znesnadnění identifikace osoby pomocí tzv. „kvazi-identifikátorů" – to jsou například datum narození, adresa a pohlaví. Pomocí kvazi-identifikátorů unikl na veřejnost např. zdravotní stav guvernéra státu Massachusetts poté, co byly zveřejněny jeho kvazi-identifikátory z volebního záznamu, které odpovídaly jeho anonymizovanému zdravotnímu záznamu. V [6] popisují autoři generalizaci dat založenou na tom principu, že nesmí existovat množiny pacientů majících stejné kvazi-identifikátory obsahující méně než K prvků. Tento přístup znesnadňuje identifikaci osob, ale byly zjištěny případy, kdy i přes K-anonymizaci byla možná identifikace. V [7] autoři vychází z K-anonymizace a navrhují zlepšení, L-anonymizaci. Ta spočívá v tom, že vyžaduje různorodost citlivých údajů v množině osob se stejnými kvazi-identifikátory.

Další možností generalizace je znepřesnění kvazi-identifikátoů. Stejná data mohou být v databázi uložena několikrát, pokaždé s různým stupněm přesnosti, s tím, že čím větší má čtenář oprávnění, k tím přesnějším datům se může dostat. Např. místo data narození se uchovává pouze rok, nebo dokonce jen dekáda. Místo adresy pouze název obce či kraje.

Jiná možnost generalizace spočívá v tom, že některé kvazi-identifikátory nebudou v databázi vůbec dostupné.

V zahraničí se pro popis chorob a symptomů používají ICD kódy (International Statistical Classification of Diseases and Related Health Problems) spadající pod správu WHO. Tyto kódy mají hierarchickou strukturu, a tudíž se přímo nabízejí ke generalizaci. Autoři [8] uvádí pravděpodobnost identifikace pacienta podle četnosti raritních ICD kódů. Doporučují odstranit 5% až 25% raritních kódů (symptomů) a nahradit je jejich generalizací. Tím se významně sníží pravděpodobnost identifikace pacienta, za relativně nízkou cenu ztráty přesnosti dat.

Netriviálnost dotazování

Velký problém týkající se výzkumu nad Big Data v biomedicínské informatice spočívá ve vlastní tvorbě dotazů. Nelze předpokládat, že by významná většina vědeckých pracovníků zabývající se biomedicínským oborem byla schopna a ochotna programovat vlastní Map/Reduce paralelní programy pro řešení dotazů nad Big Data medicínskou databází. Spíše bude pravděpodobným scénářem nastavení takové spolupráce, kdy informatici, analytici nebo programátoři vytvoří nástroje, které bude možné parametrizovat, spouštět apod.

Obě výše popsané oblasti (bezpečnost a netrivialita dotazování) by se daly vyřešit pomocí dotazovacího nástroje.

Nástroj by obsahoval "šablony" dotazů, programů nebo algoritmů, které by se daly při použití uživatelského rozhraní parametrizovat a takto použít k prohledávání databáze odbornou veřejností, aniž by musela procházet náročným procesem školení a výuky používání Big Data.

Použití šablon by jako primární cíl mělo výrazně usnadnit používání širší vědecké komunitě a zvýšit dostupnost takto zaměřeného výzkumu a jeho výstupů pro široké spektrum aplikací. Zároveň by se tímto způsobem elegantně vyřešila otázka bezpečnosti.

Příklad šablon:

- Preskripce konkrétní účinné látky dle lékařské specializace
- Analýza četnosti podle místo trvalého pobytu pacienta
- Demografická skladba pacientů
- Sezónní objemy výkonů medicínských zařízení
- Závislost výskytu chorob na druh zaměstnání pacienta

Pro tuto potřebu lze vybudovat několikavrstevnou architekturu, která bude nadstavbou základních Big Data technologií v konkrétní biomedicínské implementaci a která bude rozdělena nejméně do těchto vrstev:

1. Odborná veřejnost bude moci používat pouze předpřipravené šablony, do kterých bude možné vkládat vlastní parametry, ale nebude možné měnit podstatu dotazů. Dotaz půjde pokládat pouze jednou šablonou, šablony nelze kombinovat. Takto bude zaručeno, že nedojde k úniku citlivých údajů. Tento přístup bude sloužit např. postgraduálním studentům pro jejich základní výzkum.

2. Specializovaná pracoviště budou schopna kombinovat šablony a více parametrizovat jednotlivé dotazy. Současně však bude nad jejich činností bdít sada heuristických pravidel, jenž bude reportovat nebo ve vybraných případech i blokovat použití kombinace dotazů, které mohou přinášet riziko úniku konkretizovatelných dat.

3. Analytický tým bude připravovat šablony včetně heuristických pravidel, která budou používání takových šablon hlídat. Současně bude pod standardními bezpečnostními kontrolami (auditní log, analýza četnosti výsledků apod.) schopen Big Data sám využívat v případě, že bude existovat vysoké riziko práce

s citlivými daty. V takovém týmu budou lidé s odpovídajícím prověřením, na které budou aplikované i procesní postupy zajišťující bezpečnost využívání dat. Tento tým pak může jako součást své náplně zpracovávat komplexní úlohy a dotazy dle požadavků jednotlivých pracovišť v případech, kdy nebude efektivní použít schematické šablony a současně bude existovat riziko úniku citlivých dat. Získaná data budou před předáním žádajícímu pracovišti zkontrolována a v případě potřeby bude soubor výsledků dodatečně anonymizován.

4. Úzký specializovaný auditní tým bude nastavovat pokročilá heuristická pravidla nad celým systémem, schvalovat šablony před zveřejněním a definovat bezpečnostní pravidla.

5. Posledním prvkem může v budoucnosti být systém pracující s prvky umělé inteligence, který bude na základě souboru informací s využitím pokročilých algoritmů rozpoznávání patternů, neuronových sítí a učícího se procesu automatizovaně vyhledávat a vyhodnocovat nepodchycené principy a možnosti zneužití dat v reálném čase.

Reference

[1.] Boussi Rahmouni H, Solomonides T, Casassa Mont M, Shiu S. Modelling and Enforcing Privacy for Medical Data Disclosure across Europe. In Adlassnig KP, editor. Medical Informatics in a United and Healthy Europe – Proceedings of. Sarajevo: IOS Press ; 2009. p. 695–699.

[2.] Duncan et al. Disclosure Risk vs. Data Utility: The R-U Confidentiality map: Los Alamos National Library; 2001.

[3.] Amazon Web Services. Creating Healthcare Data Applications to Promote HIPAA and HITECH Compliance. 2012.

[4.] Alshehri , Radziszowski , Raj K. Designing a Secure Cloud-Based EHR System using Ciphertext-Policy Attribute-Based Encryption.

[5.] Narayan S, Gagné M, Reihaneh SN. Privacy preserving EHR system using attribute-based infrastructure. .

[6.] Sweeney L. k-anonymity: a model for protecting privacy. International Journal on Uncertainty. 2002; 10(5): p. 557–570.

[7.] Machanavajjhala A, Kifer D, Gehrke J, Venkitasubramaniam M. L-diversity: Privacy beyond k-anonymity. ACM Transactions on Knowledge Discovery from Data. 2007 March; 1(1).

[8.] Vinterbo S, L OM, S D. Hiding information by cell suppression. In Proc AMIA Symp; 2001. p. 726–730.

[9.] Kleinaki A, Mytis-Gkometh P, Drosatos G, Efraimidis P, Kaldoud E.: A Blockchain-Based Notarization Service for Biomedical Knowledge Retrieval. Computational and Structural Biotechnology Journal, Volume 16, 2018, Pages 288–297

[10.] Berger J., Beyr K.: Safety of Private Data in Big Data and Biomedicine. International Journal of Biomedicine and Healthcare 2015; Volume 3, Issue 1: p. 2– 5

[11.] Berger J., Beyr K.: Biomedicine data security. International Journal of Biomedicine and Healthcare 2016; Volume 4, Issue 1: p. 2–6

Kontakt

Ing. Jiří Berger, MBA
e-FRACTAL, s.r.o.
Vinohradská 1597/174
130 00 Praha 3
e-mail: *jiri.berger@e-fractal.cz*

MUDr. Jiří Kofránek, CSc.
Oddělení biokybernetiky
U nemocnice 5
128 53 Praha 2
e-mail: *kofranek@gmail.com*